

TECHNOLOGICAL CHANGE AT WORK: THE IMPACT OF EMPLOYEE INVOLVEMENT ON THE
EFFECTIVENESS OF HEALTH INFORMATION TECHNOLOGY

Adam Seth Litwin*
Carey Business School
Johns Hopkins University

January 2010

ABSTRACT

This paper uses employee and patient survey data from a large, integrated healthcare provider to assess the moderating role that employee involvement (EI) plays in the effectiveness of a patient scheduling module that is part of an electronic health record (EHR) system. The author finds that while the module facilitated the appointment-making process, its effects were greater in those clinics that sought input from frontline workers and made use of worker peers trained as system “super-users.” This study presents the first empirical evidence of EI’s potential to enhance the effectiveness of health IT, findings that should inform policymakers and sectoral actors as they allocate substantial resources toward the healthcare industry’s transition from paper-based to electronic recordkeeping. More critically, this case of workplace technological change advances work and employment theory beyond the analysis of union policies toward technological change, instead explaining how employment relations structures and processes influence the effectiveness of new technologies—IT in particular.

* The author is Assistant Professor of Management at the Johns Hopkins Carey Business School. He thanks Thomas A. Kochan, Erik Brynjolfsson, Emilio J. Castilla, and Larry W. Hunter as well as seminar participants at the Massachusetts Institute of Technology, the London School of Economics, Johns Hopkins University, and at the 2010 annual meetings of the Labor and Employment Relations Association for comments on this and earlier drafts. The project on which this paper is based won the 2009 Sloan Foundation Dissertation Prize.

Though aspects of technology and technological change have long held a key place in employment relations theory and research, the widespread diffusion of information technology (IT) over the last two decades has heightened our need to understand the workplace interplay between human and technological capital. From a policy perspective, nowhere is this need more acute than in the healthcare sector. There is near universal agreement that the industry requires major reform and that diffusion of health IT is critical to improving efficiency and service quality—a belief backed up by billions of dollars in government incentives to those adopting electronic health records (EHRs). For example, the Health Information Technology for Economic and Clinical Health (HITECH) Act allocated \$46 billion of economic stimulus funds for investment in EHR technologies. Reformers justify the allocation of these resources by blaming the slow diffusion of health IT for the poor performance of the healthcare industry, marked by skyrocketing costs and poor quality outcomes relative to other countries (Kaiser Family Foundation 2007). The Obama administration, citing a RAND Corporation study (Hillestad et al. 2005), points to a projected annual savings of \$81 billion from the effective deployment of health IT systems.

However, to date policymakers have little or no empirical evidence to support their optimistic expectations or data on the organizational or employment conditions needed to translate these new technologies into improved performance outcomes. Yet, if results from studies of investments in technologies in other industries (Batt 1999; MacDuffie and Krafcik 1992) generalize to healthcare, there is reason to question whether a “technology alone” strategy will realize policymakers’ expectations. Instead, these studies have demonstrated that technological investments need to be complemented with employment and organizational practices to achieve their desired results.

This paper begins to address this issue by testing for the performance effects of a specific type of health IT and a particular characteristic of the employment context—employee involvement (EI). It does so by examining the implementation of one piece of an EHR system, a scheduling module, across a single region of Kaiser Permanente, the nation’s largest, not-for-profit health plan. I first draw on qualitative, observational data to develop an understanding of the processes by which the scheduling module facilitates the work of frontline employees. This stage of data-gathering also allows me to identify performance measures most directly tied to the effective use of this particular technology, outcomes that are of interest to the organization itself and that are measured reliably across clinics over time. Furthermore, I

determine the specific ways in which workers and union representatives are involved in the development, deployment, and use of the IT, particularly those forms of EI that the organization believes will improve the effectiveness of the scheduling module. This qualitative evidence is then used to develop context-specific measures of the EI practices and IT in-use and to conduct a longitudinal analysis of the individual and joint effects of IT and EI on performance across multiple healthcare clinics.

The study offers a number of advantages over existing ones. It allows us to hold constant many of the unobservable contextual factors that remain unaccounted for in national, cross-industry studies of IT's performance effects (e.g., Brynjolfsson, Hitt, and Yang 2002; Caroli and van Reenan 2001). In particular, it leverages the strength of a case study approach, studying a very specific, well-defined technological change—something that cannot be done in national studies where IT is frequently defined rather vaguely (Brown and Campbell 2002). Likewise, rather than relying on measures of revenue or profit, it relies on a contextually-appropriate, homogenous performance measure as suggested by Ichniowski, Shaw, and Prennushi (1997) and MacDuffie (1995). The paper also supplements more-grounded examinations of employment practices and IT developed largely in manufacturing rather than in the service sector.

Employee Involvement and Technological Change

The past two decades have witnessed a growing body of evidence on the effects of technology and workplace practices, motivated in part by the highly visible and widely reported early experiences of General Motors (GM) and others in the auto industry with investments in automation. Case study research documented that in the 1980s, GM invested billions of dollars in automation technology—\$650 million in one GM factory alone (Kochan 1988)—but did not achieve the expected performance improvements or achieve the levels of performance observed in Japanese plants in North America or in Japan (Krafcik 1988). Instead, follow-up case study and quantitative analysis demonstrated that it was the combination of new technologies and innovative employment practices that positioned shop floor workers to “give wisdom to the machine” (MacDuffie and Krafcik 1992; MacDuffie 1995) that delivered these levels of performance. This evidence suggested that a “bundle” of innovative employment

practices, inclusive of opportunities for worker involvement in problem solving, moderated the return on investments in new technologies.¹

These results have subsequently been replicated in other manufacturing and service industries. Kelley (1996), for example, shows that increased computerization in the machined products sector drives larger productivity gains in firms that involved workers through participatory structures. More recent studies have found similar results in service and other industries. Batt (1999) found that telecommunications sales representatives with access to the new technology outperformed those not using IT, and that the size of the performance increment was greater for those workers reporting high levels of involvement in problem-solving and participation.

The most direct and most generalizable evidence of the relationship between EI and IT comes from the large sample, cross-industry econometric analyses of Brynjolfsson and colleagues (e.g., Bresnahan, Brynjolfsson, and Hitt 2000; Bresnahan, Brynjolfsson, and Hitt 2002; Brynjolfsson and Hitt 2000; Brynjolfsson, Hitt, and Yang 2002; Brynjolfsson and Hitt 2003). They show that systems of innovative employment practices seeking to involve and empower workers to make optimal use of IT are as important as the technology itself in generating performance improvements.

Just as these studies suggest that the effects of technological innovations depend on EI or related workplace practices, other studies have been unable to establish a conclusive link between EI and economic performance (Cappelli and Neumark 2001; Freeman and Kleiner 2000; Kleiner, Leonard, and Pilarski 2001). Appelbaum and Batt (1994) suggest that measurement error may be the problem, as neither researchers nor practitioners have a single, shared understanding of the meaning of EI or how it actually occurs in workplaces. Instead, the dominant finding in the literature on high-performance work systems (HPWS) has identified “bundles” or clusters of employment practices as opposed to individual practices as significant drivers of economic performance (e.g., Becker and Huselid 1998; Ichniowski, Shaw, and Prennushi 1997; MacDuffie 1995). Of course, the instrumentality of EI-inclusive bundles of employment practices also stands on firm theoretical ground. It is now widely accepted that workers will only share their valuable, often tacit, production-related information if they are

¹ Interestingly, MacDuffie’s (1995) groundbreaking empirical study measured technology very carefully in an effort to isolate the performance effects of employment practice bundles. However, it did not focus on the ways that certain employment practices managed to “unlock” new technologies.

invited to do so, have the appropriate skills to do so, and are given the appropriate incentives (Appelbaum, Bailey, Berg, and Kalleberg 2000; Becker and Huselid 1998; MacDuffie 1995).

Though the joint impact of EI on IT's performance effects has yet to be tested in a healthcare context, prior research makes clear that workplaces with high levels of EI should prove more effective at leveraging IT, particularly if one can credibly measure or hold constant those contextual variables that theory dictates be in place for EI to prove effective.

Employee Involvement and IT at Kaiser Permanente

Existing studies of the effects of EI and IT on performance have been criticized for failing to account for context (e.g., Brown and Campbell 2002; Ichniowski, Kochan, Levine, Olson, and Strauss 1996). Consequently, I describe the EHR and EI systems and the Kaiser Permanente labor management partnership (LMP) in considerable detail here in order to provide the context needed to interpret the quantitative results that follow.

Kaiser Permanente, the integrated health insurer and healthcare provider, was chosen for this study because it has been a forerunner in healthcare's conversion from paper-based to electronic recordkeeping and has a history of promoting EI as part of an overall labor management partnership (Kochan, Eaton, McKersie, and Adler 2009). Kaiser's EHR system, KP HealthConnect, once fully-deployed, will include a full complement of interoperable administrative and clinical health IT applications. One of these, which I refer to as the "scheduling module," is used for scheduling office visits, procedures, and lab tests in each region's outpatient or "ambulatory" clinics—essentially, large-scale doctors' offices.

The LMP is a cooperative arrangement between Kaiser Permanente and thirty union locals representing workers in seven of its eight regions (Kochan, Eaton, McKersie, and Adler 2009). As of 2008, the Coalition of Kaiser Permanente Unions (CKPU) and thus, the LMP, covers about 86,000 Kaiser employees. The configuration of the LMP replicates that of its management-side counterparts, creating labor-management "partners" at every level in every region in which the CKPU represents workers. At the apex of the LMP in its Oakland-based office sits a representative from Kaiser—a senior vice president reporting directly to Kaiser's COO—alongside the CKPU's director.

[—Insert Table 1 about here.—]

The LMP funds a full-time KP HealthConnect union coordinator at the national level to represent the interests of the CKPU with respect to KP HealthConnect's development, deployment, and ongoing use. It also negotiated and now administers a national-level KP HealthConnect "Effects Bargain" agreement governing job and wage protections for workers as they relate to the KP HealthConnect initiative (See Table 1.). Together, these provisions and personnel assignments establish the importance of labor to the KP HealthConnect initiative and seek to assure that KP HealthConnect will advance the interests of the workforce as it advances Kaiser's goals. Further, the agreement underlines the need for flexibility at all levels in processes and workflows and for the active involvement of labor representatives and frontline workers in developing and implementing KP HealthConnect. In exchange, the document creates and funds regional-level KP HealthConnect union representatives to represent labor alongside IT and operations leads at the top of each region's KP HealthConnect project team. Among other protections, it makes guarantees with respect to training and preparation as well as a commitment to mitigating the effects of staffing challenges that would inevitably occur in the run-up to implementation.

The Effects Bargain established the creation of at least one, full-time, KP HealthConnect labor coordinator to serve on each regional KP HealthConnect leadership team. Since the labor coordinator was charged with monitoring KP HealthConnect-related service process and workflow change experiments and pilots, he or she also assumed responsibility for identifying and responding to demands for frontline worker involvement arising in the course of the initiative. In the aggregate, Kaiser expected labor's active involvement in configuring, implementing, and eventually, encouraging optimal use of KP HealthConnect.

Kaiser's Northwest region signaled its commitment to both the Partnership and to KP HealthConnect by funding two bargaining unit employees to serve as KP HealthConnect labor coordinators, each pulled directly from the bargaining unit. With clinical functionality largely in-place, the region turned to one of KP HealthConnect's non-clinical applications, the scheduling module. The labor coordinators immediately assumed their positions on the local configuration team, alongside IT and operations leaders as well as programmers and application specialists. They also began assembling a cadre of bargaining unit members to serve as "super-users."

Super-users were support staff end-users drawn from throughout the region. At any one time, there were approximately 15-20 active super-users. They were the first to learn how

to use the scheduling module and served as liaisons between frontline support staff and the regional configuration team. As the region grew closer to implementing the system in the spring and summer of 2005, super-users were temporarily transferred on a full-time basis from their regular roles on the front lines, allowing them to travel the region answering questions and facilitating the training of other bargaining unit members. Much of what the super-users did was informal. However, there were four main channels by which their participation—and by extension, the participation of all those frontline workers whom they touched—served to make the scheduling module more effective.

First, during their travels throughout the region, they sought suggestions on how to improve the system or its rollout. Through weekly meetings, they relayed this information to the labor coordinators, who ensured it was integrated into the planning being done by the regional leadership team. It was through this process that frontline staff pointed out that the transition between scheduling systems could not be done in waves—by clinic, by department, or by any way other than what would eventually be labeled a “big bang.” This is because Kaiser patients, while assigned to a specific provider in a specific clinic, draw on services from many departments and often multiple clinics. Aside from communicating this up to management through their labor coordinators, the super-users also made a related case with respect to training, also voiced at the strategic level by the regional labor coordinators: as a consequence of the decision to go with a “big bang” rollout, all end-users would have to be trained before “go-live.”

Training was, in fact, the second area where super-users played a key role in the deployment of the scheduling module. They worked with regional trainers to develop and lead sessions for their frontline co-workers. This introductory training occurred mainly at the regional training facility, but called upon the super-users to scope out opportunities within the clinics to make sure staff were up and running on the technology. Later on in the process, they played a similar dual role in follow-up or “optimization” training.

Super-users were also charged with communicating information downward from regional leadership to those on the frontlines, a responsibility that often included as much justification as communication. For example, management’s recognition that staff from all clinics would have to be trained before the rollout reinforced the need for some extra flexibility from the rank-and-file. In particular, the short time frame meant that some training would

have to occur in the evenings and on weekends, a decision that would not be welcomed by the workforce.

Finally, super-users provided ongoing, “just-in-time” support for co-workers not only around the time of the deployment, but thereafter as well. These experts would eventually return to their jobs able to serve as their workplace’s de facto leaders and “go-to” people for all matters technological and work-related pertaining to the KP HealthConnect scheduling module. Indeed, super-users played just as vital a role in the initiative when they returned fulltime to their regular positions. Managers and frontline staff report their being in-demand as KP HealthConnect resource people in their clinics, providing co-workers with quick answers to the sorts of “just-in-time” questions that arose as those who were already formally-trained became everyday users.

Despite the sturdy structure supporting the mandate for workforce participation, interviews with frontline staff in many clinics across multiple Kaiser regions revealed a great deal of variation in just how involved workers felt they were in the project. This deviation between stated policies and their impact on the ground is actually quite common in studies linking employment practices to performance (e.g., Bartel 2004; Jones, Kalmi, and Kauhanen 2009). However, within well-defined regional boundaries, there was little or no variation in attributes of the IT module itself—including when it “went live.” Likewise, a host of contextual variables can be reasonably assumed not to vary within a single region. This study exploits these advantageous, quasi-experimental conditions.

Methods

The Technology and Organizational Setting

Through interviews with managers and labor leaders in Kaiser’s national headquarters as well as those in multiple regions, I identified an IT application, the scheduling module in Kaiser’s Northwest region, for which Kaiser management had clear and measurable performance improvement expectations. Furthermore, it was implemented in organizational units doing the same work and that were similar enough on other dimensions to provide for suitable comparisons. Headquartered in the suburbs of Portland, Oregon, Kaiser’s Northwest regional operation, Kaiser Permanente of the Northwest (KPNW), employs 880 physicians and 8,900 employees to serve just over 480,000 “members” (i.e., patients). The region spans the

greater metropolitan Portland and Vancouver, Washington areas. It offers “ambulatory” care through 27 outpatient medical office buildings, 15 of which serve as hubs for primary care—family practice, pediatrics, and internal medicine.² I focus on these primary care clinics, in part, because so many of the performance outcomes of interest to Kaiser are shaped by the member’s experience with his or her primary care physician (PCP). Bounding the sample in this way also allowed me to spend time in all of the clinics, accounting for or assuring the non-variation in contextual characteristics. For example, including appointment-making procedures beyond primary care would introduce variation across specialties and ancillary services.

The scheduling module addressed a very concrete set of organizational challenges—inefficiencies and patient dissatisfaction with the appointment-setting process. Among other challenges, those support staff charged with setting patient appointments using the legacy scheduling application frequently found themselves asking even long-term Kaiser members for data that should be permanently linked to a member’s health record number (HRN), namely contact information. The legacy system also made it difficult to schedule regularly recurring appointments and often lacked up-to-date information on providers’ availability vis-à-vis vacation scheduling, “panel support” time, or the use of planned or unplanned leave.

To understand how this would have a negative impact on economic performance, consider the process by which members make a primary care appointment by phone. They dial their clinic’s appointments line. The call is received by a member intake specialist (MIS). The MIS opens the schedule corresponding to the member’s PCP and searches for the first available appointment time or the first available time slot amenable to the member. This only disposed of about 40 percent of cases. More frequently, large sections of a provider’s schedule would be blocked as “unavailable” for one of the reasons listed above. The MIS would then transfer the member to the medical assistant (MA) supporting the appropriate provider. If the MA picked up, he or she could override or correct the schedule. If instead the MA were unavailable or serving another patient in-person, the patient calling could leave a message. If the patient ever calls again, possibly returning a call from the MA, they would start all over again at the call center, where the MIS would again try to make an appointment and would likely run into the same complication. The end result was that 75-80 percent of members initially denied an appointment would ultimately be given one within an acceptable time frame.

² The term “outpatient” is often used to describe those patients expected to check-in and out of the hospital on the same day. However, since this study does not address anything related to “inpatients” or hospital care, I use the adjectives “ambulatory” and “outpatient” interchangeably.

This chain of events came at the great expense of patient satisfaction with the appointment-making process. Furthermore, appointment-setting required 4-5 “touches” from more highly-paid MAs in addition to MISs, rather than the single touch of one MIS. Effective use of the new scheduling module was expected to address this issue and the patient dissatisfaction that arose from it.

Patient Satisfaction Survey

One reason why the scheduling module was such an attractive choice for in-depth study was its direct connection to a well-measured outcome of great interest to Kaiser managers. Whether or not the new system was effective could be measured by patients’ perceptions of the appointment-making process. As a result, Kaiser had for many years collected patient-level data on the appointment-setting process as part of a mailed paper-and-pencil Patient Satisfaction Survey sent shortly after an appointment. This study draws from two items on the Patient Satisfaction Survey to create clinic-level, monthly measures of performance. One question on the Patient Satisfaction Survey asks, “Were you able to get the appointment scheduled by talking to just one person?”. Another asks respondents to rate on a nine-point Likert-type scale their satisfaction “with the length of time spent on the phone to schedule the appointment.” These variables were strongly related. Those who answered “yes” for the binary performance item were, on average, more satisfied with the length of time required to make their the appointment ($t = 74.4, p < .000$), providing evidence of convergent validity for these performance measures (Furr and Bacharach 2008; Schwab 2005). However, the difference in discreteness allows for two, separate paths towards statistical substantiation, to be explained below.

Though the use of these types of “localized” performance measures poses a challenge for generalizability, a number of researchers have argued for their use on reliability grounds (e.g., Hunter and Pil 1995), claiming that they provide a more direct causal link than do financial performance measures. Some researchers have even chosen to use such measures even when more generalizable dollar figures could have been easily imputed (e.g., Bartel, Ichniowski, and Shaw 2007; Ichniowski, Shaw, and Prennushi 1997; MacDuffie 1995). These data from the Patient Satisfaction Survey could ultimately be drawn on to evaluate changes over time in performance measures that can be tightly tied to the scheduling module in both conceptual and practical terms. The response rate for the survey was 35%, which stacks up favorably to

comparable customer surveys administered by mail (Kaplowitz, Hadlock, and Levine 2004). Though management could not provide the necessary data dismissing the possibility of response bias, this bias should be consistent over the time period studied. Furthermore, the marketing literature suggests that disgruntled or dissatisfied patients may be more likely to respond to such surveys than others (Richins 1983). To the extent that this is true and that the use of the technology dissatisfies patients, it only serves to make the statistical estimates more conservative, i.e., biased away from theorized results.

I was able link about 43,000 patient observations to the specific PCP with whom the patient-respondent made the appointment. I could then cross these data with managerial data placing physicians into specific clinics over time and then aggregate them into a dataset of monthly observations of primary care clinics. Following Jones, Kalmi, and Kauhanen (2009) and Bartel (2004), I do this by taking the weighted means of patient responses for the physicians in that clinic in that month—a method that further strengthens the reliability of these specific performance measures (Harter, Schmidt, and Hayes 2002). For example, the dependent variable is the standardized mean of the continuous satisfaction variable for all of the Patient Satisfaction Survey responses for each clinic in each month. As a result, the point estimate on the IT “go-live” variable establishes the influence of the scheduling module on performance, whereas the estimate for a two-way interaction crossing “go-live” with a clinic’s EI score establishes the moderating impact of EI.

Survey of Support Staff

Due largely to the newness of the technology, Kaiser’s human resource (HR) records did not contain reliable measures of EI. Though Kaiser conducts an annual poll of its employees, the instrument had only recently been augmented with a single and very broad question about the health IT system. Therefore, I administered an employee survey that achieved an overall response rate of 63%—388 MAs and MISs who are expected to use the technology in the course of their everyday work. Analyses confirmed that those MISs who responded had about the same average age and job tenure as those who did not. The MA responders had the same average tenures as their non-responding colleagues. However, those MAs who responded were

marginally older, on average, than those that did not respond—41.8 years vs. 39.3 years ($t = 2.44, p < .01$).³

There are a number of advantages to surveying employees directly, and then aggregating these data to the clinic (i.e., establishment) level. First, EI measures cannot be biased by individual, clinic-level managers wanting to offer an idealized account of EI (Jones, Kalmi, and Kauhanen 2009). Second, Huber and Power (1985) suggest that single-response bias be tackled by asking survey questions of the “most-informed respondent” in the establishment. This design goes a step further by drawing on multiple, informed respondents from each clinic. That is, only those employees expected to use the scheduling module in the course of their everyday work are included in the analysis. This technique also avoids “frame of reference” problems (Hunter and Pil 1995) by asking frontline workers the very EI-related questions that they should know the answers to—not questions about a broad EI construct. Furthermore, Gerhart, Wright, MacMahan, and Snell (2000) suggest that drawing on multiple respondents from each establishment disposes of inter-rater reliability issues, though they also note that research designs bounded to a small number of clinics and a homogenous group of workers rarely suffer from this problem anyway. Finally, perhaps the most significant methodological challenge to studies linking employment practice “inputs” to performance “outputs” occurs when the same instrument is used to collect both. In this way, so-called “common method bias” generates artificially-inflated correlations between EI, IT use, and performance (Podsakoff, MacKenzie, Lee, and Podsakoff 2003). However, the research design circumvents the causes of common method bias with its collection of the independent and dependent variables from completely separate and unrelated sources—one long in existence for organizational use and another conceived of and administered years later purely for the purposes of this research.

The employee survey included questions derived from my observations and interviews to measure EI relevant to employees in this particular organizational setting, similar to the approach adopted by Bidwell (2009). According to Jarvis, Mackenzie, and Podsakoff (2003), the construction of formative indicators such as these rather than more traditional “reflective” measures makes sense when indicators “define” different aspects or dimensions of the construct and when indicators need not be interchangeable. In Kaiser’s case, there were multiple ways in

³ I could not test for randomness with respect to sex. However, nearly all of the MAs and MISs sampled were women.

which workers might have participated in the IT initiative, and any one of them could effectively substitute for any other. For example, a worker may have been directly canvassed for their thoughts on effective system use. Alternatively, they may have relied frequently on guidance from a super-user. Summing answers into a composite measure therefore captures the overall level of EI in this context, even though there is no *a priori* reason to expect a high correlation between items (Bidwell 2009).⁴

Responses to eight survey items were summed to construct the EI index. The first four, items are: 1.) My suggestions relating to the design and improvement of [the scheduling module] have been valued., 2.) My issues or complaints about it have been ignored., 3.) There is at least one bargaining unit member in my office who helps me be a better user of [the scheduling module], and 4.) Before it was rolled out, the people whose work would be affected were asked for guidance. Each was answered using a seven-point, Likert-type scale in which 7 represented strong agreement. The second item was reverse-coded. The remaining four items were binary in nature. Respondents answered questions on whether or not a fellow member of the bargaining unit introduced them to the scheduling module, provided them with their follow-up training on the module, or otherwise served as an on-site expert or “super-user” for the scheduling module. Respondents also answered yes or no as to whether they provided any specific recommendations on additional ways that the system could be used to meet its strategic goals. As a further reliability check, the survey included an open-ended question asking workers to document a specific suggestion that they had made. This step provided additional confidence that respondents understood exactly the kinds of EI they were being asked about (Hunter and Pil 1995).

Estimation Strategies

I employ a two-prong approach to show that the use of the IT is associated with performance increases at the clinic level and that these effects are greater in those clinics with higher mean levels of EI. First, I run a series of nested, multilevel regression models on a data set of clinic-months. The time-constant nature of the EI measure unfortunately precludes the estimation of clinic-level fixed effects with these data. Therefore, I take a second, separate approach to substantiate the performance impact of the scheduling module as well as the

⁴ This scale proves only marginally reliable by conventional standards ($\alpha = .58$). Nonetheless, a low alpha does not indicate low reliability in the case of formative measures like that employed for EI (Bidwell 2009; Bollen and Lennox 1991; Jarvis, Mackenzie, and Podsakoff 2003).

moderating role that EI played in increasing the effectiveness of the new technology. I run 14 separate logistic regressions—one for each clinic—using patient-level responses to the Patient Satisfaction Survey as the unit-of-observation.⁵ In this case, I use the other performance measure, whether or not a patient was able to schedule their appointment with the first person he or she spoke to. These regressions include just four independent variables—two time trends and two IT-related dummies. As with the previous set of regression models, the first IT dummy captures the transformation period into the new system, and the second captures the performance impact of IT “go-live.”

Results

Table 2 presents summary statistics from the survey of the Northwest’s support staff. Means are calculated using only responses from those MAs and MISs expected to use the scheduling module in the course of their work. The first set of variables represents the four continuous items contributing to the EI scale. Notice how in all four cases, means hover near the neutral response (4 = “neither agree nor disagree”), albeit with significant variation about the mean. Overall, only 11 percent of respondents claimed that they were first introduced to the technology by a fellow member of the bargaining unit (as opposed to a manager or an IT staffer), though 18 percent asserted that they had, in fact, received follow-up training from a co-worker. About 40 percent noted the importance of “super-users”—fellow members of the bargaining unit pulled from their regular, frontline positions to assist in the development and deployment of the system—to their successful use of the scheduling module. Interestingly, about 15 percent of respondents have made specific recommendations of ways that the system could be used more effectively, the details of which I could validate with the responses to a free-form text field included in the survey. For example, some workers suggested the need for “write” privileges in addition to “read-only” privileges at certain screens. Others pointed out the need to make sure that a patient’s contact details remain on-screen throughout the appointment-setting process or the need to allow the home phone number field to be left empty for those patients having only a cell phone. Others had suggested the creation of shortcuts for frequently-used “bundles” of mouse clicks, like those required to make certain, regularly-occurring types of office visit appointments.

⁵ While there are 16 clinics under study, the regressions include only 14 or 15 clinics, depending on the particular model. This is because the Peterson clinic closed prior to the collection of quantitative measures of EI, and the Ulrich clinic opened too late in the observation period to provide pre-“go-live” observations.

[—Insert Table 2 about here.—]

Table 3 breaks out the dependent variables for each (de-identified and relabeled) clinic, derived from patient-level data. The first three columns focus on the binary dependent variable—whether or not the patient was able to make an appointment with the first person he or she spoke to on the telephone. In the Bruford clinic, for example, of 3,911 patient responses to the question over the observation period (October 2004 to June 2007), 78 percent answered affirmatively. Note that most of the clinics average around 80 percent for this variable over the period of observation. The one exception appears to be Collins, which only managed to schedule appointments with one “touch” 73 percent of the time. The next three columns repeat the exercise for the continuous dependent variable—patient’s satisfaction with the length of the phone call required to make the appointment. In this case, the variable was standardized such that the mean was equal to zero and the standard deviation equal to one. Therefore, each clinic’s mean for the variable as reported in Table 3 is relative to the overall sample average. The Fleetwood clinic averaged .2 standard deviations above the sample mean, the highest of all the clinics. The clinic labeled Mullen achieved the lowest performance and the widest variation on this metric over the sample period.

[—Insert Table 3 about here.—]

Table 4 displays the first set of models estimated, beginning with a simple model considering only the effects of a linear time trend. The first model shows a small, but statistically significant month-to-month increase in the dependent variable between October 2004 and June 2007. Once a separate, post-implementation trend is added on the right-hand side (in the second model), the estimated partial slope on the original time trend turns negative and remains so for the remaining models to be estimated. The “go-live” time trend, however, that first appears in the second model reveals a positive association between the use of the scheduling module and the performance measure it was intended to influence. Despite the negative, month-to-month effect of trend, the post-implementation time trend is actually positive and remains so for all subsequent estimates. Consistent with anecdotal accounts, customer service was suffering prior to the implementation of the scheduling module, a trend

that reversed itself at the same time as the transition to the new system. The next model adds two dummy variables capturing transition to and deployment of the scheduling module. Both estimates are positive and statistically significant in this and the remaining models. Also note the point estimate on the post “go-live” time trend doubles. That means that once we account for a structural break in the time series, we can see evidence of a large (.44 standard deviations), one-time jump in performance as well as a steady, sizable (.06 standard deviations) month-to-month performance increase associated with the scheduling module, despite what would otherwise be a declining performance function (-.05 standard deviations each month) over time. These effects are not sensitive to changes in the way the transition period is operationalized, e.g., one month or two months on either side of the transition from legacy systems to the new IT.

[—Insert Table 4 about here.—]

The last two models in Table 4 incorporate the effects of EI on the effectiveness of the technology. Model 4 incorporates only a main effect for EI. Interestingly, this predictor has an estimated performance effect that is insignificantly different from zero, suggesting that the impact of EI comes not through an engaged workforce *per se*, but from the moderating impact of EI with respect to the scheduling module initiative. It is also worth noting that the inclusion of the EI variable in the fourth model does virtually nothing to the point estimates of all those variables carried over from the three versions of the equation previously estimated. The fifth and final model in Table 4 adds the two-way interaction to directly capture the incremental, moderating effect of EI on the IT-performance link. Controlling for all of the other effects, an increase of one standard deviation in the EI index increases the effectiveness of the technology by .27 standard deviations. Interestingly, the estimate for the main EI measure turns negative, further demonstrating that EI’s performance impact appears to come through its moderation of the scheduling module’s effect on performance. Furthermore, the results are robust to many different ways of operationalizing the EI measure.

Figure 1 projects the point estimates for IT “go-live” for each clinic—derived from the 14 separate logit estimates—on a scatterplot as a function each clinic’s mean EI score. First, notice that accounting for trend and transition, none of the clinics witnessed a performance decrement arising from the technology, and three of them increased their performance by at

least one standard deviation. More important, the figure reveals a positive association between workers' involvement in the effort and the size of performance improvements: those clinics whose workers reported greater mean EI levels with respect to the technology saw greater performance improvements arising from the technology than those clinics with lower mean EI scores. Once again, the shape of the point cloud—revealing the positive association between the size of a clinic's IT-engendered performance gain and its mean level of EI—is robust to different operationalizations of EI.

[—Insert Figure 1 about here.—]

Discussion and Conclusions

This study of how Kaiser Permanente used EI to achieve larger returns from newly-deployed IT sheds light on the moderating role that aspects of the employment relationship play in linking technology to performance. It does so by examining the same clinics before and after the technology was turned on. In particular, Kaiser's deployment of its scheduling module, one component of its much larger EHR system, was associated with clinic-level performance improvements. However, these improvements were more than 50% greater in those clinics in which workers scored more highly on a contextual measure of EI. It appears that while the scheduling module provided workers across all the clinics additional, real-time information on provider availability and patient information, employees made better use of that information when they understood management's strategic rationale for the system, when they were able to communicate their own ideas and concerns back up to the strategic level, and when they were aided fellow frontline workers who could ease them through the deployment process.

Though employment relations has long acknowledged the role of technology in its theory-building (e.g., Dunlop 1958 [1993]; Slichter 1941; Slichter, Healy, and Livernash 1960), much of that focus has been on trade union responses to new technologies that were intended to serve as substitutes for labor. Inasmuch as this study illustrates the ways that a union-management-negotiated agreement and participatory employment practices enable workers to use IT more effectively, it focuses explicitly on the ways that employment practices influence the effectiveness of new technologies intended to make workers more productive.

Furthermore, it does so outside of manufacturing, the sector that has been the focus of most of the empirical work to date on the employment practice correlates of organizational performance. Therefore, not only do the results inform the fastest growing sector of the US economy—healthcare—they offer a lens into the service sector and “service processes” (as opposed to manufacturing’s “production processes”) more broadly.

The design of the study also allows for a clean separation of the technology inputs from the EI inputs that management theory suggests complement one another in production. The great benefit of IT is that it makes more information available to frontline workers (Bresnahan, Brynjolfsson, and Hitt 2002; Brynjolfsson and Mendelson 1993). However, pushing information downward and outward—in this case, up-to-date information on patients and on physician availability—will do much less to influence performance if those workers who will need to use the technology cannot shape how it is used and are not “brought on board” with clear communication from managers and union representatives. For example, with respect to the scheduling module at Kaiser, aspects of the technology and of the organization necessitated that some training had to occur outside of regular working hours, and the labor coordinators and the super-users played a key role in justifying this unpopular decision to the region’s workforce.

Aside from shedding some much-needed light on IT and EI in the service sector, the results also offer an explanation for the ambiguous relationship between EI and organizational performance in received studies. Recall from the last two columns of Table 4 that the direct effects of EI on performance were insignificantly different from zero. Those positive performance effects resulting from EI appeared to arise entirely through the interaction of the new technology and the EI measure. Aside from this study, research in work and employment has yet to explore this possibility explicitly, largely because of the discipline’s tendency to view technology, at best, as a “black box” that needs to be measured or controlled for, but not necessarily examined in-depth. This oversight is particularly interesting given the aforementioned attention that industrial relations theory has long given to the “technological context” for employment relations. As IT and other new technologies become even more ubiquitous, employment relations scholars would do well to look within both the IT and EI processes at work to better understand how and why they interact to affect performance outcomes.

The immediate implication for both policymakers and healthcare administrators is that health IT can improve organizational outcomes. Therefore, it makes sense that the government should promote the diffusion of EHRs and related technologies, and it makes sense for practices to respond accordingly to those incentives. However, policies that seek only to encourage the adoption of health IT as opposed to the adoption of both the technology and the employment practices that more-fully “unlock” it are, at best, incomplete. Such costly mandates—like those that appear in the 2009 stimulus package—should also include language to encourage the adoption of employment structures and processes along the lines of those employed at Kaiser.

The results of this study might be challenged on a number of grounds, namely the overall reliability and validity of either the EI or performance measures, econometric endogeneity, common method bias, or recall bias on the part of employees. The issues of reliability and construct validity are the most critical, in part because both the EI and performance measures were developed or chosen specifically for this study rather than taken from previously validated instruments. As noted above, this study was guided by the principle that internal validity and context-specific measures should be privileged over generalizability. However, decisions on how to measure EI and performance were made only after spending time in the field. The resulting EI measures, while not broadly generalizable, were those that emerged as relevant to making this specific technology work in this specific context. Likewise, the performance measures, both tied to the appointment-making process, while not generalizable to other IT applications, were ones that mattered to management and that were conceptually linked to the effective use of the technology. This intensive and context specific information should provide some confidence in the validity of the measures. In the cases of both EI and performance, I have drawn on multiple measures that reveal a consistent relationship between IT and EI “inputs” and performance “outputs.” These results should raise one’s confidence in the reliability of these measures.

With respect to endogeneity, the quantitative analyses do not explicitly control for confounding factors at the clinic level. The standard, fixed-effects approach to dealing with this issue was not available because the measure of EI did not vary over time. As a result, the two-way interaction of interest could not have been estimated. A related set of issues associated with the estimates—those supporting Figure 1 as well as those displayed in Table 4—is that of omitted variable bias. In particular, one might argue that those clinics that were “ready” for the

technology based on observed measures of EI or some other unobserved factors, not surprisingly, were able to use the technology more effectively. With respect to these issues, reliance on qualitative investigation in addition to the statistical estimates offers some assurance of the findings' overall validity. For example, it was through the deliberative, pre-statistical investigative process that I determined that the “go-live” date was set at the regional level and was not chosen clinic-by-clinic based on each clinic's readiness.

Regarding common method bias, this issue would be of greater concern had I collected the performance measures from workers themselves, particularly had I used the same instrument to measure technology and employment relations variables. However, common method bias should be less of a concern here since the performance measures come not from workers, but from patients, using an instrument that had been developed prior to this study and for more general purposes than to evaluate the effectiveness of new technologies (Podsakoff, MacKenzie, Lee, and Podsakoff 2003). It is also the case that the EI measures were specific to the IT initiative under study as opposed to being indicative of the larger employment relations context.

With respect to both common method and potential recall bias, it is worth noting that the main effect of EI in the IT initiative turns negative—though not at conventional levels of statistical significance—in the final regression estimate in Table 4, the one that includes the critical, two-way interaction between EI and the use of the scheduling module. This implies that after controlling for trend and for the use of the technology, those clinics with more-engaged workforces performed relatively poorly over the observation period, counter to the idea that solid performance led workers to evaluate the employment relations climate more positively than they would have otherwise. In the final estimate, the positive effects of EI are channeled entirely through the use of the new system, an effect that would be difficult to explain as the product of endogeneity, common method bias, or recall bias on the part of employee respondents.

Finally, given the unique features of the Kaiser labor management partnership, further work is needed to determine if similar effects are observed in more traditional, unionized settings and/or in nonunion settings that provide other employee voice arrangements. Given also that these are still “early days” for this industry's much-needed technological innovation, this area begs for greater study from scholars of work and employment.

REFERENCES

- Appelbaum, Eileen, Thomas Bailey, Peter Berg, and Arne Kalleberg. 2000. *Manufacturing Advantage: Why High-Performance Work Systems Pay Off*. Ithaca: ILR Press.
- Appelbaum, Eileen, and Rosemary L. Batt. 1994. *The New American Workplace: Transforming Work Systems in the United States*. Ithaca: ILR Press.
- Bartel, Ann P. 2004. "Human Resource Management and Organizational Performance: Evidence from Retail Banking." *Industrial and Labor Relations Review*, Vol. 57, No. 2, pp. 181-203.
- Bartel, Ann, Casey Ichniowski, and Kathryn Shaw. 2007. "How Does Information Technology Really Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement and Worker Skills." *Quarterly Journal of Economics*, Vol. 122, No. 4, pp. 1721-1758.
- Batt, Rosemary L. 1999. "Work Organization, Technology, and Performance in Customer Service and Sales." *Industrial and Labor Relations Review*, Vol. 52, No. 4, pp. 539-564.
- Becker, Brian E., and Huselid, Mark A. 1998. "High Performance Work Systems and Firm Performance: A Synthesis of Research and Managerial Implications." In Gerald Ferris, ed., *Research in Personnel and Human Resources Management*. Greenwich, CT: JAI Press, pp. 53-101.
- Bidwell, Matthew. 2009. "Do Peripheral Workers Do Peripheral Work? Comparing the Use of Highly-Skilled Contractors and Regular Employees." *Industrial and Labor Relations Review*, Vol. 62, No. 2, pp. 200-225.
- Bollen, Kenneth, and Richard Lennox. 1991. "Conventional Wisdom on Measurement: A Structural Equation Perspective." *Psychological Bulletin*, Vol. 110, No. 2, pp. 305-314.
- Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt. 2002. "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm Level Evidence." *Quarterly Journal of Economics*, Vol. 117, No. 1, pp. 339-376.
- Bresnahan, Timothy F., Brynjolfsson, Erik, and Hitt, Lorin M. 2000. "Technology, Organization, and the Demand for Skilled Labor." In Margaret M. Blair and Thomas A. Kochan, eds., *The New Relationship: Human Capital in the American Corporation*. Washington, DC: Brookings, pp. 145-184.
- Brown, Clair, and Benjamin A. Campbell. 2002. "The Impact of Technological Change on Work and Wages." *Industrial Relations*, Vol. 41, pp. 1-33.
- Brynjolfsson, Erik, and Lorin M. Hitt. 2003. "Computing Productivity: Firm-Level Evidence." *Review of Economics and Statistics*, Vol. 85, No. 4, pp. 793-808.

Brynjolfsson, Erik, and Lorin M. Hitt. 2000. "Beyond Computation: Information Technology, Organizational Transformation and Business Performance." *Journal of Economic Perspectives*, Vol. 14, No. 4, pp. 23-48.

Brynjolfsson, Erik, Lorin M. Hitt, and Shinkyu Yang. 2002. "Intangible Assets: Computers and Organizational Capital." *Brookings Papers on Economic Activity*, Vol. 2002, No. 1, pp. 137-181.

Brynjolfsson, Erik, and Haim Mendelson. 1993. "Information Systems and the Organization of Modern Enterprise." *Journal of Organizational Computing*, Vol. 3, No. 3, pp. 245-255.

Cappelli, Peter, and David Neumark. 2001. "Do 'High-Performance' Work Practices Improve Establishment-Level Outcomes?" *Industrial and Labor Relations Review*, Vol. 54, No. 4, pp. 737-775.

Caroli, Eve, and John van Reenan. 2001. "Skill-Biased Organizational Change? Evidence from a Panel of British and French Establishments." *Quarterly Journal of Economics*, Vol. 116, No. 4, pp. 1449-1492.

Dunlop, John T. 1958 [1993]. *Industrial Relations Systems*. Boston: Harvard Business School.

Freeman, Richard B., and Morris M. Kleiner. 2000. "Who Benefits Most from Employee Involvement: Firms or Workers?" *American Economic Review Papers and Proceedings*, Vol. 90, No. 2, pp. 219-223.

Furr, R. Michael, and Verne R. Bacharach. 2008. *Psychometrics: An Introduction*. Thousand Oaks, CA: Sage.

Gerhart, Barry, Patrick M. Wright, Gary C. MacMahan, and Scott A. Snell. 2000. "Measurement Error in Research on Human Resources and Firm Performance: How Much Error Is There and How Does It Influence Effect Size Estimates?" *Personnel Psychology*, Vol. 53, No. 4, pp. 803-834.

Harter, James K., Frank L. Schmidt, and Theodore L. Hayes. 2002. "Business-Unit-Level Relationship Between Employee Satisfaction, Employee Engagement, and Business Outcomes: A Meta-Analysis." *Journal of Applied Psychology*, Vol. 87, No. 2, pp. 268-279.

Hillestad, Richard, James Bigelow, Anthony Bower, Federico Girosi, Robin Meili, Richard Scoville, and Roger Taylor. 2005. "Can Electronic Medical Record Systems Transform Health Care?: Potential Health Benefits, Savings, and Costs." *Health Affairs*, Vol. 24, No. 5, pp. 1103-1117.

Huber, George P., and Daniel J. Power. 1985. "Retrospective Reports of Strategic-Level Managers: Guidelines for Increasing Their Accuracy." *Strategic Management Journal*, Vol. 6, No. 2, pp. 171-180.

- Hunter, Larry W., and Pil, Frits K. 1995. "How do You Survey Firms?" In Paula B. Voos, ed., *Proceedings of the 47th Annual Meeting of the Industrial Relations Research Association*. Madison: IRRA, pp. 152-162.
- Ichniowski, Casey, Thomas A. Kochan, David Levine, Craig Olson, and George Strauss. 1996. "What Works at Work: Overview and Assessment." *Industrial Relations*, Vol. 35, No. 3, pp. 299-333.
- Ichniowski, Casey, Kathryn Shaw, and Giovanna Prennushi. 1997. "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines." *American Economic Review*, Vol. 87, No. 3, pp. 291-313.
- Jarvis, Cheryl Burke, Scott B. Mackenzie, and Philip M. Podsakoff. 2003. "A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research." *Journal of Consumer Research*, Vol. 30, No. 2, pp. 199-218.
- Jones, Derek C., Panu Kalmi, and Antti Kauhanen. 2009. "How Does Employee Involvement Stack Up? The Effects of Human Resource Management Policies in a Retail Firm." *Industrial Relations*, Vol. 49, No. 1, pp. 1-21.
- Kaiser Family Foundation. 2007. *Employee Health Benefits: 2007 Annual Survey*. Washington, DC: KFF.
- Kaplowitz, Michael D., Timothy D. Hadlock, and Ralph Levine. 2004. "A Comparison of Web and Mail Survey Response Rates." *Public Opinion Quarterly*, Vol. 68, No. 1, pp. 94.
- Kelley, Maryellen R. 1996. "Participative Bureaucracy and Productivity in the Machined Products Sector." *Industrial Relations*, Vol. 35, No. 3, pp. 374-399.
- Kleiner, Morris M., Jonathan S. Leonard, and Adam M. Pilarski. 2001. "How Industrial Relations Affects Plant Performance: The Case of Commercial Aircraft Manufacturing." *Industrial and Labor Relations Review*, Vol. 55, pp. 195.
- Kochan, Thomas A. 1988. "On the Human Side of Technology." *ICL Technical Journal*, Vol. 6, No. 2, pp. 391-400.
- Kochan, Thomas A., Adrienne E. Eaton, Robert B. McKersie, and Paul S. Adler. 2009. *Healing Together: The Labor-Management Partnership at Kaiser Permanente*. Ithaca: Cornell.
- Krafcik, J. F. (1988). Comparative analysis of performance indicators at world auto assembly plants. Unpublished master's thesis, MIT/Sloan School.
- MacDuffie, John Paul. 1995. "Human-Resource Bundles and Manufacturing Performance: Organizational Logic and Flexible Production Systems in the World Auto Industry." *Industrial and Labor Relations Review*, Vol. 48, No. 2, pp. 197-221.

MacDuffie, John Paul, and Krafcik, John F. 1992. "Integrating Technology and Human Resources for High-Performance Manufacturing: Evidence from the International Auto Industry." In Thomas A. Kochan and Michael Useem, eds., *Transforming Organizations*. New York: Oxford University Press, pp. 209-226.

Podsakoff, Philip M., Scott B. MacKenzie, Jeong-Yeon Lee, and Nathan P. Podsakoff. 2003. "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies." *Journal of Applied Psychology*, Vol. 88, No. 5, pp. 879-903.

Richins, Marsha L. 1983. "Negative Word-of-Mouth by Dissatisfied Consumers: A Pilot Study." *The Journal of Marketing*, pp. 68-78.

Schwab, Donald P. 2005. *Research Methods for Organizational Studies*. Mahwah, NJ: Lawrence Erlbaum Associates.

Slichter, Sumner H. 1941. *Union Policies and Industrial Management*. Washington, DC: Brookings Institution.

Slichter, Sumner H., James Joseph Healy, and E. Robert Livernash. 1960. *The Impact of Collective Bargaining on Management*. Washington, DC: Brookings Institution.